

Jani Björklund

KUVAN KORJAAMINEN ILMAN SEN NÄKEMISTÄ

Semanttisen informaation hyödyntäminen kuvan
korjauksessa

Informaatioteknologian ja viestinnän tiedekunta
Kandidaattitutkielma
Toukokuu 2019

TIIVISTELMÄ

Jani Björklund: Kuvan korjaaminen ilman sen näkemistä
Kandidaattitutkielma
Tampereen yliopisto
Sähkötekniikka
Toukokuu 2019

Kamerateknologian kehittyessä kameroilla voi saada yhä tarkempia ja realistisen näköisiä kuvia riippumatta millainen kamera on. Erilaiset algoritmit kykenevät uudelleenrakentamaan valokuvan, vaikka olisi vähemmän informaatiota alkuperäisestä kuvasta tarjolla.

Yksi kameroiden suurimmista ongelmista on värikonstanssin korjaaminen kuvissa. Erilaiset valaistukset ja heijastukset saavat valokuvaan oudon sävyn, mikäli kuvan värikonstanssia ei korjata algoritmilla. Nykypäivänä yhä useampi yritys on kehittänyt omia algoritmejaan, sekä tekoälyjä ratkaisemaan tätä ongelmaa kameroissa.

Tässä työssä tutkitaan sitä, että voisiko kuvan värikonstanssin korjaamiseen hyödyntää semanttista informaatiota ympäristöstä. Työssä kategorioidaan Shi-Gehlerin kuvakokoelmaa pienempiin alakategorioihin, joista etsimme yhtäläisyyksiä. Tällä tiedolla voimme tietää voisiko kuvan värikonstanssin korjata tietämällä itse kuvan.

Avainsanat: Värikonstanssi, Shi-Gehler, Kamera

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

ABSTRACT

Jani Björklund: Fixing an image without seeing it
Bachelor's thesis
Tampere University
Electrical Engineering
May 2019

Camera technology has been improved in many ways. Nowadays the cameras do not need the latest hardware to take realistic looking photos. Different kind of algorithms can create the photos from less information.

One of the biggest problems in cameras has been fixing the color constancy. Different kind of lightings and reflections can form unnatural colors to the photo, if the photo is not processed with an algorithm. Nowadays companies are building to fix color constancy in the photos.

This thesis studies if a photo can be fixed utilizing semantic information of the surroundings. Shi-Gehler's dataset is categorized into subcategories to find out if it is possible to get decent looking photos knowing only what is in the photo. If there is a connection between those subcategory photos, we can fix images using this assumption.

Keywords: Color constancy, Shi-Gehler, Camera

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

ALKUSANAT

Tämän työn myötä vajaa kolme vuotta kestäneen kandidaatintutkinnon opinnot saadaan pakettiin, ja ensimmäinen virstanpylväs diplomi-insinöörin tutkintoa kohti on saavutettu. Kandidaatintyön merkitys on minulle henkilökohtaisesti suuri, sillä sain näytettyä itselleni pystyväni tämän suoritukseen.

Tätä en olisi saavuttanut ilman erityisiä ihmisiä ympärilläni. Haluan kiittää ohjaajaani Joni Kämäräistä tämän aiheen löytämisestä, sekä tukemisesta ja innostamisesta. Vertaisteni lisäksi Sähkökillassa haluan myös kiittää tämän lukuvuoden aikana elämässäni erittäin oleellisessa osassa ollutta Growthmakers-kurssin opiskelijoita. Heidän ansiostaan henkinen jaksamiseni riitti tämänkin työn valmistumiseen.

Seuraavaksi määränpäänäni on saavuttaa diplomi-insinöörin tutkinto. Tästä kokemuksesta on erittäin paljon hyötyä sen tavoittelemisessa.

Tampereella, 16. toukokuuta 2019

Jani Björklund

SISÄLLYSLUETTELO

1	Johdanto	1
2	Teoria	3
2.1	Värikonstanssi	3
2.2	Ihmisen näköjärjestelmä	3
2.3	Kameran toiminta	4
3	Tutkimusongelma	6
4	Semanttinen värikonstanssi	8
4.1	Shi-Gehler kuvakokoelma	8
4.2	Testausympäristö	8
4.3	Testaus	9
4.3.1	Ulkokuvat	10
4.3.2	Sisäkuvat	12
4.4	Lopputulos	14
5	Yhteenveto	15
	Lähdeluettelo	16

KUVALUETTELO

1.1	Sama valokuva neljässä eri valaistuksessa. Värikonstanssia ei ole korjattu [3]	1
2.1	Silmän rakenne. [7]	4
2.2	Kameran rakenne. [8]	4
2.3	Esimerkki sensorin rakenteesta. Jokainen väri kuvaa yhtä pikseliä, ja mitä aallonpituutta kyseinen osa hyväksyy. [9]	5
3.1	Esimerkki värikonstanssilla korjaamattomasta ja korjatusta kuvasta [6]	6
4.1	Kaavio tuloksista. Arvoina kuvien määrä kategoriassa ja alle kolmeen asteeseen jäävien osuus tästä määrästä	14

TAULUKKOLUETTELO

4.1	Keskiarvot kaikista värikanavista	9
4.2	Ulkona otetut kuvat	9
4.3	Sisällä otetut kuvat	9
4.4	Ulkokuvien tulokset	11
4.5	Sisäkuvien tulokset	13
4.6	Lopputulos semanttiselle mallille	14

LYHENTEET JA MERKINNÄT

L_{src}	Valaisevan valon väri
S_{src}	Heijastavan esineen reflektanssi
λ	Aallonpituus
$C(k, sen)$	Kameran sensorin herkkyys
RGB	Värimalli, joka pohjautuu punaiseen, vihreään ja siniseen väriin)

1 JOHDANTO

Jo vuosikymmenien ajan ihmiset ovat halunneet ikuistaa hetken muistoksi ottamalla siitä valokuvan. Sen takia kamera kännykässä tuntuu itsestäänselvyydeltä, jolla pystytään ottamaan kuvan hieman spontaanimmastakin tilanteesta. Nykyisin ei tarvitse enää taidemaalaria ikuistamaan hetkeä lähes realistisesti, kun voidaan kaivaa taskusta vain pienen kämmenen kokoisen laitteen ja saada sekuntien murto-osassa aikaan parempaa jälkeä kuin käsin maalattuna. Harva nuoremmista ihmisistä osaisi kuvitella elämää ilman nyky-päivän sosiaalista mediaa kuvineen.

Vaikka kamerateknologia on saanut alkunsa jo satoja vuosia sitten, kun vuonna 1826 *Nicéphore Niépce* otti ensimmäisen valokuvan valoherkälle paperille, niin värien vääristymistä ei ole saatu estettyä kätevästi valokuvauksen aikana. Tämä saa kuvalle muodostettua väärienlaisen sävyn, joka nähdään erittäin häiritseväenä kuvaa tarkasteltaessa.

Värikonstanssi on ilmiö, jossa ihmisen näkö kykenee tunnistamaan esineen värin riippumatta siihen kohdistuvan valaistuksen väristä. Ihmisen aivot kykenevät päättämään, millainen valaistus on kyseessä ja täten muuttamaan sen prosessoidessaan muodostuvaa kuvaa vastaamaan valkoisessa valossa otettua kuvaa. [3] Vastaavanlaista valon värin tunnistusta ei löydy vielä kameroista, mikä vaikeuttaa kuvan värin korjaamista. Tästä johtuen kamera joutuu arvaamaan minkä värinen valo on kyseessä.



Kuva 1.1. Sama valokuva neljässä eri valaistuksessa. Värikonstanssia ei ole korjattu [3]

Ongelman ratkaisemiseksi eri yritykset ovat kehittäneet ohjelmistotasolla algoritmeja, jotka korjaavat vääristymisen tasapainottamalla RGB-kanavien värejä lähemmäksi valkoista valoa vastaavaksi. Tällöin kuva muistuttaa enemmän ihmisen näkemää. Tekoälyn yleisyydessä myös värikonstanssin korjaamista varten on kehitelty oppivia ohjelmistoja korjaa-

maan entistä paremmin tätä ongelmaa.

Tässä työssä tutkitaan värikonstanssia itsessään ilmiönä sekä mistä tämän ilmiön aiheuttama ongelma valokuvaamisessa johtaa juurensa. Lopuksi vielä tutkitaan, voisiko tätä ongelmaa ratkaista hyödyntämällä semanttista informaatiota, eli tietoa kuvattavasta ympäristöstä. Tämän mallin toimintaa tutkitaan empiirisesti hyödyntämällä Shi-Gehlerin kuvakokoelmaa.

2 TEORIA

Värikonstanssi on yksi kamerateknologian keskeisistä ongelmista. Tavoitteena on kuvan vangitsemisen digitaaliseen muotoon luomalla todellisuutta vastaavan kuvan värityksen. Ymmärtääkseen miten kamera kykenee muodostamaan oikean näköisen valokuvan näkemästään, tulee ymmärtää värikonstanssin lisäksi, miten ihmisen näkö toimii ja miten kamera mallintaa näköä muodostamalla realistisen näköisen kuvan.

2.1 Värikonstanssi

Värikonstanssin toi yleiseen tietoisuuteen kemisti *Michel Eugene Chevreul* vuonna 1839, kun hän oivalsi värien vuorovaikutusilmiön ja sen, että värit pysyvät esineessä itsessään vakiona. Ainoa asia mikä muuttuu, on esineeseen kohdistuva valaistus. Erilainen valaistus saa värit näyttämään eri sävyisiltä, mutta esineen ominaisväri pysyy samana.[5] Tämän takia esimerkiksi kaksi samanväristä paperia vaikuttavat eri sävyisiltä, kun toisen asettaa varjoon ja toisen suoraan valaistuksen alle.

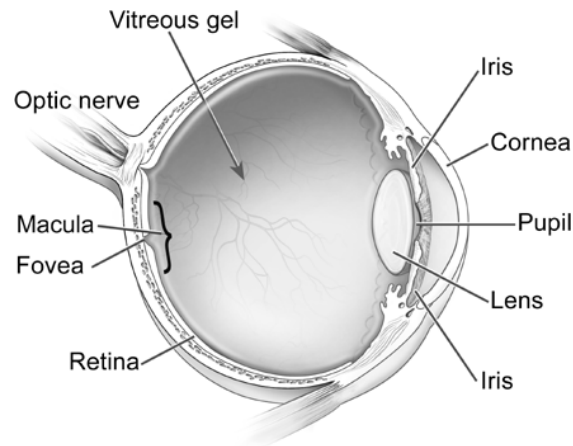
Valo itsessään voi heijastua monesta eri suunnasta, minkä takia sitä on erittäin vaikeaa mallintaa luonnossa. Tämän lisäksi valonlähteitä voi olla useampi, kuten huoneessa olevat lamput ja ikkunasta paistava auringonvalo.

2.2 Ihmisen näköjärjestelmä

Ihmisen näkö perustuu silmässä useamman osan yhteistyöhön. Ihmisen näkö kykenee muuntamaan silmään saapuvan valon aallonpituudet aivoille prosessoitavaksi signaaliksi.

Valo läpäisee aluksi silmän uloimman osan sarveiskalvon (*engl. cornea*). Sarveiskalvo on silmän uloimpana osana oleva kaareva kudos, jonka tehtävänä on suojata silmää, ja taittaa valoa läpi iirishimentimen (*engl. iris*) aivan mykiöön (*engl. lens*) asti. Iirishimentimen tarkoituksena on himmentää valon määrää silmään. Tämän läpipäästetty valon mykiö kohdistaa silmän perällä olevalle verkkokalvolle (*engl. retina*).

Verkkokalvolla sijaitsee miljoonia tappisoluja (*engl. cone cell*) ja sauvasoluja (*engl. rod cell*), jotka muuntavat saapuvan valon yhteistyössä valoinformaatioksi. Kyseinen valoinformaatio kulkeutuu näköhermoja pitkin aina aivoihin asti, jossa aivot muodostavat näistä signaaleista ”kuvan”. Koska värit itsessään ovat vain eri aallonpituuksia, tappi- ja sau-

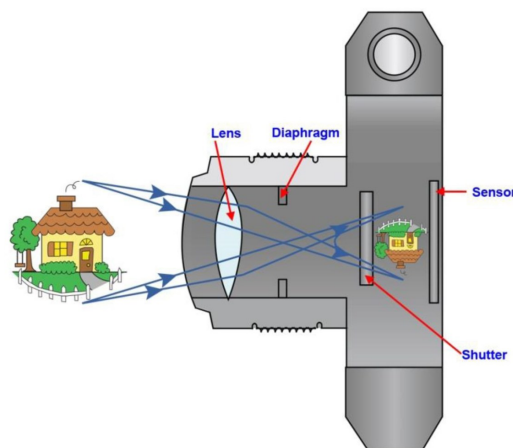


Kuva 2.1. Silmän rakenne. [7]

vasolut kykenevät tunnistamaan värejä erottelemalla aallonpituuksia toisistaan. Tämän takia osa sauvasoluista on herkkiä punaiselle, osa siniselle ja osa vihreälle alueelle näkyvän valon spektrissä (350 – 750nm). [4]

2.3 Kameran toiminta

Kamera toimii hyvin samalla periaatteella kuin silmä kuvan ottamisessa. Kuten silmässäkin, valo matkaa linssin (*engl. lens*) läpi kameran perältä löytyvään sensoriin (*engl. sensor*), joka käsittelee saapuvan valon muodostaen kuvan. Eri kameroilla on erilaisia komponentteja ja tapoja käsitellä saapuvaa valoa, mutta nyt keskitytään vain yksinkertaisiin digitaalisiin kameroihin.



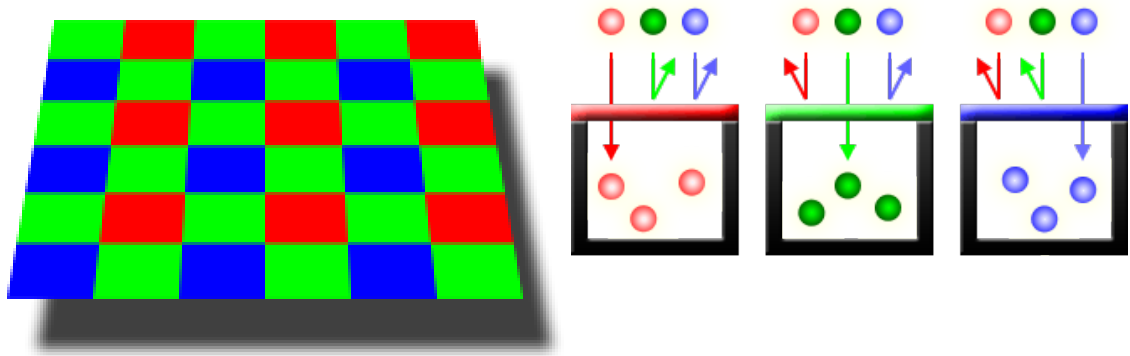
Kuva 2.2. Kameran rakenne. [8]

Kamerassakin valo kulkee ulkoisen linssin läpi, kuten myös silmässä. Kyseisessä linssissä valo heijastuu kameran perällä olevaan sensoriin. Linssin heijastumista voidaan hienosäätää linssistä mahdollisesti löytyvistä optiikoista. Polttoväliä ja täten myös zoomausta kaukaisempiin kohteisiin voidaan kameroissa muuttaa zoomaavilla linseillä, jolloin polttoväliä kameran sisällä linssin ja kameran perällä olevan sensorin välillä kasvatetaan.

[8]

lirishimmennintä silmässä vastaa valotusaukko (*engl. diaphragm*). Iirishimmennimen taivoin, valotusaukko kontrolloi paljonko valoa päästetään läpi kameraan. Yksinkertaisimmissa kameroissa ei kumminkaan löydy tämän kaltaista ominaisuutta, jotenka valoa tulee aina yhtä paljon läpi.

Silmän toiminnassa jäi käsittelemättä silmäluomi, joka kontrolloi milloin valoa päästetään läpi silmään asti. Kamerassa samanlaista tehtävää hoitaa kameran suljin (*engl. shutter*), joka vaikuttaa valolle annettuun aikaan eli valotusaikaan. Valotusajalla tarkoitetaan kauanko kamera ottaa vastaan valoa kuvaa ottaessaan. Sillä on merkitystä erityisesti silloin, jos kuvataan pimeämmässä ympäristössä. Pimeämmässä ympäristössä valoa ei ole yhtä paljon, jolloin kuvaan tarvittavan informaation saanti olisi huonompaa. Pimeämpi ympäristö vaatii pidemmän valotusajan, mutta myös vakaamman kuvausalan, jotta vältetään tärähdyksistä johtuvista epäterävyyksistä. Kunnolla valaistussa ympäristössä valotusaika on vain sekunnin murto-osia, eli esimerkiksi vain 1/250 sekuntia.



Kuva 2.3. Esimerkki sensorin rakenteesta. Jokainen väri kuvaa yhtä pikseliä, ja mitä aallonpituutta kyseinen osa hyväksyy. [9]

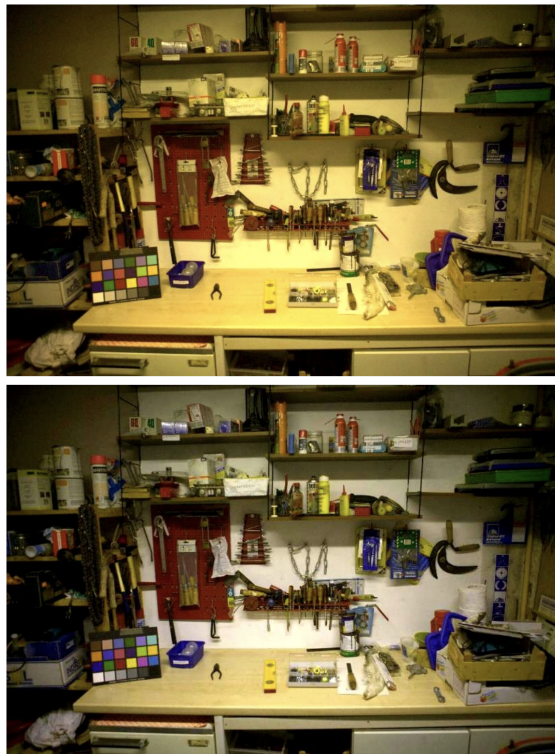
Lopulta valo saapuu kameran perällä olevalle sensorille. Sensori on jaettu pienen pieniin ruutuihin, joista jokainen keskittyy yhteen kolmesta väristä [Kuva 2.3]. Tällöin noin puolet pisteistä on vihreitä värejä, neljäsosa punaisia ja loput sinisiä. Koska vain osa väreistä otetaan huomioon, niin puuttuvat värit saadaan interpoloimalla näistä väreistä. Näistä pisteistä pystytään rakentamaan lopulta kuva. [9]

3 TUTKIMUSONGELMA

Kuten edellisen luvun lopussa mainittiin, jokainen kameran sensorin osa päästää vain tiettyä väriä lävitseen, mistä myöhemmin muodostuu kuva. Tämä arvo lasketaan seuraavalla kaavalla:

$$I_k(x, y) = \int_{\lambda} L_{src}(\lambda) S_{ref}(x, y, \lambda) C_{k, sen}(\lambda) d\lambda \quad (3.1)$$

Kaavassa 3.1 L_{src} kuvaa kohdetta valaisevan valon väriä, S_{src} heijastavan esineen reflektanssi ja $C_{k, sen}$ kameran sensorin herkkyyttä. Näistä kolmesta muuttujasta muodostetaan pikselin arvo kuvassa omalla värikanavallaan. Puuttuvat värikanavat pikselissä muodostetaan erilaisia interpolointitapoja hyödyntäen.



Kuva 3.1. Esimerkki värikonstanssilla korjaamattomasta ja korjatusta kuvasta [6]

Korjaamattomassa kuvassa törmätään eroavaisuuteen ihmisen näkemään verrattaessa. Ihmisen näkösystemi kykenee tunnistamaan valon värin L_{src} , jonka takia valo tulee aivoille saapuvassa kuvassa käsiteltyä vastaamaan valkoista valoa. Tämän takia ihminen ei näe ympäristöään valon värin vaikutuksen alaisena.

Sen sijaan tätä valaistuksen korjausta ei automaattisesti tehdä kamerassa, joka vaikuttaa merkittävästi siihen minkä värisenä kamera näkee esineet. Ilman värikonstanssia korjavia algoritmeja kamerassa, kuvaan yleensä jää valon väriä vastaava sävy. Esimerkissä [Kuva 3.1] nähdään, miten kuvan sävy on ilman korjausta selkeästi keltaisempi kuin mitä se todellisuudessa olisi.

Luonnollisen kuvan muodostamiseksi pitää saadut pikselien arvot muuttaa vastaamaan valkoisessa valossa otettua kuvan arvoja. Tämä tapahtuisi yksinkertaisesti muuttamalla yhtälön [Yhtälö 3.1] valon väri vastaamaan valkoista väriä. Tässä ongelmana kumminkin on se, että kameral on vaikeaa erottaa kuvatessa valaisevan valon väriä ja heijastavan esineen reflektanssia. Tätä vaikeuttaa entisestään se, että valo pystyy heijastumaan monesta pinnasta ja valon lähteitä voi olla useampi.

Kuvan värikonstanssin korjaamiseen on kehitelty monenlaisia algoritmeja. Pääpiirteittäin ne voi jakaa kahteen kategoriaan: ennalta kalibroituviin ja dataan pohjautuviin algoritmiin. Ennalta kalibroidut algoritmit olivat yleisempiä 1980-luvulla, ja niissä käytetään ennalta laskelmoitua tietoa valokuvattavista esineistä kuvan korjaamiseen. Tämä ei ole kumminkaan kannattavaa nykypäivänä valokuvauksen tapahtuessa vaihtelevissa paikoissa. Dataan pohjautuvat algoritmit pohjautuvat pitkälti koneoppimiseen ja sensoridatan käsitteilyyn. Esimerkiksi *Gray World* -algoritmissa otetaan värikanavien keskiarvot, jolla korjamalla voidaan saada tasapainotettu värikonstanssi. Tämän algoritmin toiminta perustuu oletukseen, että kuvan värien keskiarvo on väriltään harmaa. [1]

Kuvan korjaamiseen on myös nyt tuotu tekoälyyn pohjautuvia ratkaisuja. Google AI:n tuella on muun muassa toteutettu *Fast Fourier Color Constancy* niminen algoritmi, jonka suorituskyky on jo yksi parhaimmista mitä tällä hetkellä löytyy [2].

4 SEMANTTINEN VÄRIKONSTANSSI

Selvittääksemme semanttisen informaation merkitystä kuvien värikonstanssin korjaamisessa, tarvitsemme korjaamattomia kuvia, sekä niitä varten tarkoitettuja korjausarvoja jokaiselle RGB-arvolle. Tätä varten on valmis kuvakokoelma nimeltään Shi-Gehler kuvakokoelma [10].

4.1 Shi-Gehler kuvakokoelma

Shi-Gehlerin kuvakokoelma on *Peter Gehlerin* johtaman tutkimusryhmän käsittelemien kuvien kuvasarja, joka koostuu 568 kuvasta eri paikoista Englannissa [6]. Kyseiset kuvat ollaan käsitelty uudestaan *Lilong Shin* ja *Brian Funtin* toimesta, jonka seurauksena kuvat ovat saatu vähemmän prosessoituun muotoon, jotta saadaan parempilaatuisia kuvia tutkimukseen [10]. Kuvakokoelma koostuvat Canon 5D ja Canon 1D kameroilla otetuista kuvista. Kaikista kuvista löytyy värikalibrointiin käytettävä taulu, jolla pystytään kalibroimaan arvot oikeaksi. Kyseisiin kuviin on jätetty värikonstanssi korjaamatta.

Värikontrastilla korjaamattomien kuvien lisäksi kokoelmasta löytyy myös jokaista kuvaa vastaavat korjausarvot RGB-kanaville. Näillä kolmella arvolla käsittelemällä kuvan voidaan saada kuvasta värikontrasti korjattua näyttämään taas luonnolliselta. Tämä kokoelma siis sopii täydellisesti testaamaan semanttisen informaation avulla korjaamista.

4.2 Testausympäristö

Semanttisen informaation testausta lähdetään suorittamaan jakamalla jokaisen kuvan yhä pienempiin ja pienempiin kategorioihin informaation perusteella. Esimerkkeinä voisi esimerkiksi olla kuvat ulkona ja sisällä, sekä sisätiloissa kuvat vessassa ja keittiössä. Jokaisen kategorian korjausarvot lasketaan keskiarvallisesti yhteen, jotta saadaan jokaisesta kategoriasta vastaavat kolme RGB-arvoa.

$$Error = \arccos \frac{\hat{v}_1^{-1} \cdot \hat{v}_2}{\|\hat{v}_1\| \|\hat{v}_2\|} \quad (4.1)$$

Koska käytännössä vertaillaan eroavaisuutta kolmesta oikeasta RGB korjausarvosta koostuvaa vektoria ja kolmesta lasketusta korjausarvosta koostuvaa vektoria, niin toimiva tapa

virheen laskemiseen on kulmavirhe [Kaava 4.1]. Kulmavirheessä tarkastellaan kolmiulotteisten vektoreiden välistä kulmaa. Mitä lähempänä arvo on nollaa, sitä lähempänä vektorit ovat toisiaan.

Sanotaan, että ihminen ei kykene erottamaan vääränlaista värikonstanssia kuvassa, mikäli ero todellisiin RGB arvoihin on alle kolme astetta. Tästä johtuen testaus pohjautuu tutkimaan sitä, kuinka moni korjaus menee alle kolmeen asteeseen todellisesta arvosta. Näiden lisäksi jokaisesta alakategoriasta otetaan minimi-, mediaani ja keskiarvot ylös.

4.3 Testaus

Tutkitaan millainen alkutilanne olisi, mikäli muodostetaan kaikista yksi RGB korjaus kaikista kuvien korjausarvoista keskiarvot ottamalla. Vähemmän yllättävästi saadaan jo virhettä aikaiseksi.

Taulukko 4.1. Keskiarvot kaikista värikanavista

Pienin virhe	Mediaanivirhe	Keskivirhe	Alle kolme astetta %	Kokonaismäärä
7.77	17.75	16.65	0.00%	568 kpl

Taulukosta [Taulukko 4.1] huomataan, ettei yksikään kuva pääse alle kolmen asteen virheeseen. Pieninkin arvo on itsessään kaksinkertainen määritellyyn kolmen asteen virheeseen verrattaessa.

Seuraava vaihe on kategorioida kuvia alakategorioihin, jotta saadaan semanttiselle informaatiolle vastaavat kuvasetit. Ensimmäinen jakaminen on suurimpiin pääkategorioihin: ulkona otettuihin kuviin [Taulukko 4.2], sekä sisällä otettuihin kuviin [Taulukko 4.3]. Tämän jakamisen jälkeen tulokset olivat tällaiset.

Taulukko 4.2. Ulkona otetut kuvat

Pienin virhe	Mediaanivirhe	Keskivirhe	Alle kolme astetta %	Kokonaismäärä
0.23	2.12	1.51	80.06%	316 kpl

Taulukko 4.3. Sisällä otetut kuvat

Pienin virhe	Mediaanivirhe	Keskivirhe	Alle kolme astetta %	Kokonaismäärä
1.53	9.54	9.63	2.78%	252 kpl

Tästä huomataan jo erittäin suuri ero sisällä ja pihalla otettujen kuvien välillä. Ulkona otetuista kuvista [Taulukko 4.2] yli neljä viidestä kuvasta saataisiin pelkästään tällä jaolla korjattua riittävän hyvin. Sen sijaan sisällä otetuissa kuvissa [Taulukko 4.3] vain muutama prosentti onnistuisi riittävän hyvin. Tälle todennäköisimpänä syynä voidaan pitää valaistuksen vaihtelua sisätiloissa, kun taas ulkona otetuille kuville pääasiassa voidaan pitää auringonvaloa ainoana valaistuksena.

4.3.1 Ulkokuvat

Seuraavaksi jaetaan ulkona otetut kuvat alempiin kategorioihin. Ulkokuvien isoimpina kategorioina ovat selkeästi kuvat rakennuksien läheisyydessä ja kuvat täysin luonnossa. Testimielessä jaetaan kuvat rakennuksellisiin ja kuviin ilman rakennuksia. Nämä ja kaikki muut ulkokuvien tulokset sijoitetaan taulukkoon [Taulukko 4.4].

Tästä jaosta huomataan mielenkiintoisesti, että rakennuksellisilla kuvilla on hieman parempi suoritus verrattaessa kuviin, mistä ei ole havaittavissa rakennuksia. Rakennuksettomissa kuvissa metsien varjo muuttaa hieman valaistusta ja täten myös valaistusta, joka on todennäköisimpänä syynä aavistuksen isompiin virheasteisiin.

Rakennuksia on sijainniltaan kahdenlaisia kuvakokoelmassa. On kuvia, jotka oli otettu kaupungissa, ja kuvia, jotka oli otettu taajamassa. Näiden perusteella pystytään jakamaan kuvat kaupunkikuviin ja taajamakuviin.

Kaupungissa otettujen kuvien suoriutuminen jatkoi suurin piirtein samalla tasolla rakennuskuvien kanssa, eikä mitään suurempaa mullistusta tapahtunut. Taajamakuvien tapauksessa korjaus parani aavistuksen, sillä esiintynyt virhe pieneni. Pienin virhe kasvoi, mutta mediaanivirhe ja keskivirhe laskivat aavistuksen.

Koska taajamakuvien määrä alkoi olla jo melko pieni, niin kategoriointia jatketaan vain kaupunkikuvien suhteen ulkokuvien haarassa. Testimielessä koitetaan kahdenlaista jakoa ja niiden suoriutumista: jakoa ihmisten esiintymisen mukaan ja jakoa autojen esiintymisen mukaan.

Arvoja tarkastelemalla voidaan huomata ihmisten suhteen jaon nostavan keskiarvoa. Virhe pysyy alle kolmessa asteessa päälle 4/5 kuvista, mutta tulos ei muutu kovin paljoa muuten. Pienin virhe laskee verrattaessa kaupunkikuvien jakoon. Seuraavaksi selvitetään miten autolliset kuvat suoriutuvat samaisesta tehtävästä.

Autojen tarkastelu suoriutui erittäin huonosti verratessa ihmisillä tunnistamiseen. Ensimmäistä kertaa ulkona otetuissa kuvissa päädyttiin alle 80% onnistumiseen, kun tarkasteltiin autoja kaupungissa. Tämä tulos oli osittain yllättävä, sillä kuvittelisi auton pinnasta johtuvista heijastuksista olevan vaikutusta heijastuksiin liittyen.

Tässä tilanteessa voidaan todeta, että kyseistä haaraa ei ole järkeä jatkaa eteenpäin tulosten huonontuessa. Ihmisten suhteen kategoriointi tuottaa parempaa tulosta, joten päädytään käyttämään sitä autojen sijasta. Valitaan seuraavaksi haara, jossa todettiin, ettei kuvissa ole rakennuksia.

Rakennuksettomissa kuvissa on vaikeaa löytää suurempia kategorioita. Yksi isompi ero on: kuvat, joissa on pääasiassa luontoa, ja kuvat, joista löytyy esimerkiksi teitä.

Jako kuvien välillä onnistui lähes tasaisesti. Valitettavasti kuvien korjaaminen ei antanut lupaavia tuloksia. Keskimäärin kummankin mediaanivirhe oli lähempänä kolmea astetta kuin aikaisemmissa tarkastelluissa tapauksissa. Tästä on vaikeaa jatkaa alemmas, paitsi

koittamalla uudestaan asfaltillisissa kuissa ajoneuvojen vaikutusta. Mielenkiinnosta tarkastellaan tuloksia, joita saadaan jakamalla kuvat ajoneuvojen esiintymisen mukaan.

Autoilla jakamisesta huomataan, ettei päästä paljon parempiin tuloksiin edellisestä jaosta. Autottomissa kuissa voidaan huomata pientä arvojen paranemista, joten jako ei ollut täysin hyödytön. Autollisissa kuissa osittain ongelmana oli pieni kuvien määrä, joka saattaa vaikuttaa myös mediaanivirheen suuruuteen.

Tämän alemmaksi meneminen on käytännössä mahdotonta nykyisillä jaoilla, joten kerätään tulokset yhteen taulukkoon vielä. Kaikki ulkokuvien tulokset on esitetty taulukossa 4.4.

Taulukko 4.4. Ulkokuvien tulokset

<i>Pienin virhe</i>	<i>Mediaanivirhe</i>	<i>Keskivirhe</i>	<i>Alle kolme astetta %</i>	<i>Kokonaismäärä</i>
Ulkokuvien keskiarvo				
0.23	2.12	1.51	80.06%	316 kpl
Kuvissa on rakennuksia				
0.23	1.83	1.29	83.33%	198 kpl
Kuvissa ei ole rakennuksia				
0.35	2.59	2.10	77.12%	118 kpl
Kuvat on otettu kaupungissa				
0.24	1.94	1.39	82.73%	139 kpl
Kuvat on otettu taajamassa				
0.30	1.60	1.16	84.75%	59 kpl
Kuvissa on ihmisiä kaupungissa				
0.11	1.87	1.65	82.69%	52 kpl
Kuvissa on kaupunkia ilman ihmisiä				
0.20	1.95	1.40	82.76%	87 kpl
Kuvissa on autoja kaupungissa				
0.37	2.07	1.78	70.59%	17 kpl
Kuvissa on kaupunkia ilman autoja				
0.09	1.88	1.23	84.30%	121 kpl
Kuvat on otettu täysin luonnossa				
0.73	2.73	2.47	65.46%	55 kpl
Kuvista löytyy asfalttia luonnosta				
0.12	2.52	1.98	79.37%	63 kpl
Kuvista löytyy autoja luonnossa				
0.31	2.98	1.79	72.73%	22 kpl
Kuvat otettu luonnossa ilman autoja				
0.03	2.28	1.93	80.49%	41 kpl
Ulkokuvien lopullinen tulos				
0.03	2.12	1.52	79.05%	316 kpl

Ulkokuvien lopullisessa tuloksessa huomataan tuloksen huonontuneen prosenttiyksikön

verran ja virhe pysyy lähes muuttumattomana. Ainoastaan pienin virhe pieneni lähes olemattomaksi, mutta se on laiha lohtu tässä tilanteessa. Voisi sanoa ettei ulkokuvien käsittelyllä ole merkitystä kovin paljoa.

4.3.2 Sisäkuvat

Valaistuksen suhteen on olemassa monenlaisia lamppuja, jotka hehkuvat eri aallonpituuksia ja näin myös värejä. Sisätiloissa valaistus myös saattaa olla hieman himmeämpi, joten kuvat itsessään ovat pimeämpiä. Tämä yhdistettynä vielä ikkunan läpi tulevaan valoon saa tilanteesta erittäin monimutkaisen värinkonstanssin kannalta.

Sisällä otetut kuvat on hieman vaikeampaa jakaa täsmällisesti pienempiin kategorioihin, sillä kuvaaja on kuvannut melko pitkälti työpaikallaan, kotonaan sekä yksittäisissä rakennuksissa kaupungissa, kuten ostoskeskuksessa ja kirkossa. Kaupungilla otetuissa kuvissa on niin pieniä kokonaisuuksia, että tämä vaikuttaisi erittäin vahvasti tulokseen vääristävästi.

Ensimmäinen jako sisällä otettujen kuvien suhteen on kokeilla jakaa kuvat työpaikkakuviin ja kotona otettuihin kuviin. Kumpaankaan kategoriaan menemättömät kuvat eritellään kategoriaan ”muut”. Tämän perusteluna on idea tutkia miten toimistoissa yleisesti tavattava valaistus vaikuttaa verrattaessa esimerkiksi kotona oleviin valoihin. Tulokset laitetaan muiden sisäkuvien tavoin taulukkoon [Taulukko 4.5].

Hieman yllättäen muut kategoria selvisi paljon paremmin verrattaessa muihin sisätilojen kategoriointeihin. Tulokset ovat toki huonoja, koska alle viidenneksen kuvista korjautuisi tällä logiikalla. Työpaikan kuvien laajuus oli arvattava ongelma, sillä kuvia oli otettu käytävillä, toimistoissa ja kokoustiloissa.

Seuraavaksi tarkastellaan, miten tulos parantuisi, mikäli työpaikkakuvat jaettaisiin esimerkiksi toimiston ja muiden toimitilojen välillä. Muut toimitilat sisältävät tässä tilanteessa esimerkiksi taukotilat ja käytävät.

Työpaikkajaosta voidaan huomata, että muiden toimitilojen valaistuksen vaihtelevuus vaikuttaa aika merkittävästi kuvien värikonstanssiin. Kun taas siirrytään toimistoon, niin kuvan korjaus paranee hieman. Nyt on saavutettu hieman yli viidenneksen korjattavissa olevista kuvista.

Toimistotilojen suhteen koitetaan uudestaan, mikäli ihmisten oleminen kuvassa tuo jotain merkitystä kuvien suoritukseen. Tämän perusteella voidaan todeta, mikäli on kannattavaa ihmisiä edes tunnistaa kuvista.

Tästä kokeilusta voidaan todeta, ettei ihmisiä kannata tunnistaa kuvasta, sillä se ei vaikuta juuri ollenkaan tulokseen positiivisella tavalla. Ainoastaan yksi kuva käytännössä menee alle kolmen asteen.

Viimeiseksi tarkastellaan vielä kotona otettuja kuvia tarkemmin. Onko vaikutusta sillä,

että kuvasta löytyy huonetta valaiseva ikkuna? Kategorioimme kotikuvat ikkunallisiin ja ikkunattomiin kuviin.

Ikkunan suhteen jakamisesta voi huomata, ettei kannata kyseistä jakoa suorittaa, sillä tulos huonontui hieman entisestään. Ikkunallisista kuvista vain yksi meni alle kolmen as-teen, ja ikkunattomista vain kaksi.

Lopputulos kumminkin parantui hieman verrattaessa alkuperäiseen jakoon. Sisällä ole-vien kuvien haaran lopputulos on seuraava 4.5:

Taulukko 4.5. Sisäkuvien tulokset

<i>Pienin virhe</i>	<i>Mediaanivirhe</i>	<i>Keskivirhe</i>	<i>Alle kolme astetta %</i>	<i>Kokonaismäärä</i>
Sisäkuvien keskiarvo				
1.53	9.54	9.63	2.78%	252 kpl
Kuvat on otettu kotona				
0.95	11.79	10.49	9.23%	65 kpl
Kuvat on otettu työpaikalla				
2.32	7.39	7.48	2.42%	124 kpl
Kuvat on otettu muualla				
1.07	7.36	5.59	17.46%	63 kpl
Kuvissa on toimistotilat				
1.11	6.61	5.67	22.50%	40 kpl
Kuvissa on muut toimitilat				
1.66	7.39	6.70	5.95%	84 kpl
Kuvissa on ihmisiä toimistossa				
2.63	6.61	5.40	4.76%	21 kpl
Kuvissa on toimistotilat ilman ihmisiä				
2.05	6.17	5.40	4.76%	21 kpl
Kuvat otettu ikkunattomassa kodissa				
0.75	11.51	9.88	7.55%	53 kpl
Kuvat on otettu ikkunallisessa kodissa				
2.11	11.36	6.99	8.33%	12 kpl
Sisäkuvien lopullinen tulos				
0.94	8.30	6.87	9.43%	252 kpl

Sisäkuvien lopullisesta tulokseta huomataan, että ollaan saavutettu noin kahdeksan pro-sentin nousu tulokseen. Tämä tulos jää kumminkin alle kymmenekseen kaikista sisäku-vista, joka ei vaikuta kovin hyvältä tulokselta. Tarkemmalla valaistukseen liittyvällä tarkas-telulla olisi voinut mahdollisesti saada parannettua tulosta merkittävämmiin.

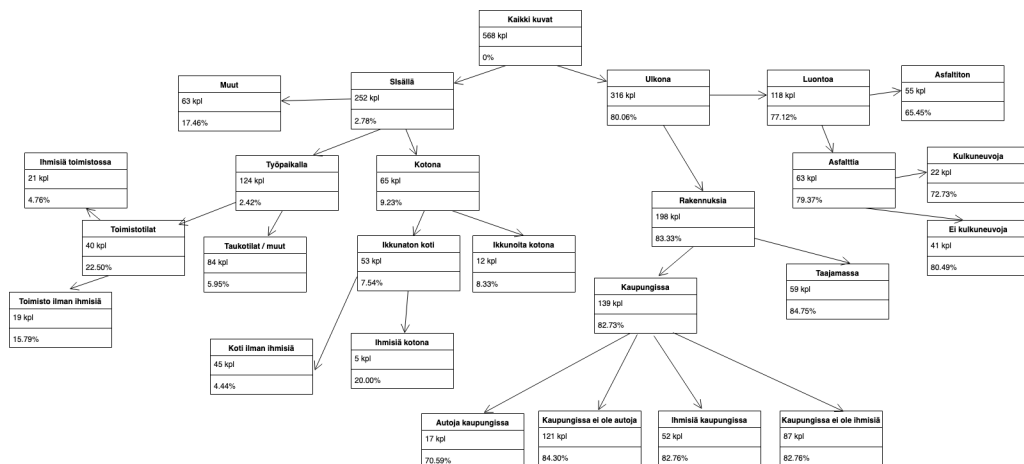
4.4 Lopputulos

Yhdistämällä edeltävien kohtien tulokset, voidaan tarkastella kuinka semanttiseen informaatioon tukeutuva malli suoriutui. Yhdistämällä alakategorioiden arvot saadaan lopputulokseksi:

Taulukko 4.6. Lopputulos semanttiselle mallille

Pienin virhe	Mediaanivirhe	Keskivirhe	Alle kolme astetta %	Kokonaismäärä
0.03	4.82	3.16	48.66%	568 kpl

Tulokset eivät ole kovin lupaavia kokonaisuudessaan. Vajaa puolet kuvista saavutti toivotun tason. Vaikka pienimmillään päästiin 0.03 asteen suuruiseen virheeseen, niin tämäkin on selitettävissä ulkona olevan valaistuksen yksipuolisuuteen.



Kuva 4.1. Kaavio tuloksista. Arvoina kuvien määrä kategoriassa ja alle kolmeen asteen jäävien osuus tästä määrästä

Kuten kaaviosta [Kaavio 4.1] huomataan, aika monessa kategoriassa ollaan päästy pie-
niin kuvamääriin. Harmiksi muut-kategorioissa ei ollut kovin paljon yhtenäisyyksiä, jotta
niistä olisi voinut jatkaa jakamista eteenpäin mahdollisen tuloksen paranemisen kannalta.

5 YHTEENVETO

Kuvien jakamisella saavutettiin pientä paranemista, mitä pidemmälle mennään kategorioinnissa. Ulkona otettujen kuvien taso säilyi koko tutkimuksen korkeana, mutta sisällä otettujen kuvien tulos parani aavistuksen kategorioinnin myötä.

Lopputuloks oli pettymys sisällä otettujen kuvien suhteen, jonka syytä on vaikeaa päätellä pelkästään katsomalla kuvia. Teoriassa sisätiloissa on monenlaista valaistusta erilaisten lamppujen ansiosta. Tästä johtuen myös lamppujen väreissäkin on suuria eroavaisuuksia. Tämän lisäksi ulkoa tuleva auringonvalo ikkunoiden läpi vaikuttaa valon väriin merkittävästi.

Ulkona otettujen kuvien tulos yllätti erittäin suuresti, ja antoi hieman uskoa onnistuvaan tutkimukseen. Ulkona otetuissa kuvissa tulokset ovat itsessään niin hyviä, että melkein tällä matalalla tasolla, millä toimittiin, voitaisiin jo hyödyntää oikeisiin kuviin. Tämä tosin on selitettävissä pelkän auringonvalon valaistuksella, joka yksinkertaisti korjausta merkittävästi.

Kokonaisuudessaan mallin suoriutuminen oli sen verran huonoa, että jatkotutkimista kannattaa harkita erittäin paljon. Isommalla kuvakokonaisuudella tai tarkemmalla kategorioinnilla saatetaan päästä lupaavimpiin tuloksiin. Voidaan kumminkin miettiä, onko tämä malli kannattava verrattaessa tehokkaimpiin kilpaileviin malleihin. Tulevat kandidaatintyön kirjoittajat ja muut tutkijat voivat pohtia asiaa itse.

LÄHDELUETTELO

- [1] B. R. Abidi, A. Koschan ja M. A. Abidi. An Overview of Color Constancy Algorithms. Tutkielma. Purdue University, maaliskuu 2006.
- [2] J. T. Barron ja Y.-T. Tsai. *Fast Fourier Color Constancy*. <https://storage.googleapis.com/pub-tools-public-publication-data/pdf/4b134951.pdf>. (Viitattu 05.01.2019).
- [3] A. Baslamisli, A. Gijsenij ja T. Gevers. http://colorconstancy.com/?page_id=9. (Viitattu 05.01.2019).
- [4] M. F. Bear, B. W. Connors ja M. A. Paradiso. *Neuroscience: Exploring the Brain, 3rd Edition*. Lippincott Williams ja Wilkins, 1996.
- [5] M. E. Chevreul. *The Principles of Harmony and Contrast of Colour and Their Applications to the Arts*. Schiffer Pub Ltd, 1987.
- [6] P. Gehler, C. Rother, A. Blake, T. Minka ja T. Sharp. Bayesian Color Constancy Revisited. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2008.
- [7] N. E. Institute. *How we see*. <https://nei.nih.gov/healthyeyes/howweseesee>. (Viitattu 05.01.2019).
- [8] A. Kandimalla. *How does a digital camera work?* <https://lightartacademy.com/blog/basics/how-does-a-digital-camera-work/>. 5. marraskuuta 2014. (Viitattu 05.01.2019).
- [9] S. McHugh. *Cambridge in Colour*. <https://www.cambridgeincolour.com/tutorials/camera-sensors.htm>. (Viitattu 03.01.2019).
- [10] L. Shi ja B. Funt. *Re-processed Version of the Gehler Color Constancy Dataset of 568 Images*. <http://www.cs.sfu.ca/~colour/data/>. (Viitattu 05.01.2019).